

University of California

UC Linked Data Project Team Report

Systemwide Work, Opportunities, and Recommendations

FINAL REPORT

October 1, 2018

Arwen Hutt (Chair), UC San Diego

Kevin Balster, UC Los Angeles

Noah Geraci, UC Riverside

Rachel Jaffe, UC Santa Cruz

Haiqing Lin, UC Berkeley

Chrissy Rissmeyer, UC Santa Barbara

Carl G Stahmer, UC Davis

Kathryn Stine, California Digital Library

Table of Contents

| | |
|--|-----------|
| Executive Summary | 2 |
| Background | 3 |
| Linked Data Use Cases | 4 |
| University of California Libraries Linked Data Activity | 8 |
| External Linked Data Activity | 10 |
| Systemwide Opportunities | 12 |
| Recommendations | 16 |
| Conclusion | 19 |
| Appendix 1: Snapshot of UC Libraries Linked Data Projects & Activities | 20 |
| Appendix 2: UC Libraries Linked Data Survey | 25 |
| Appendix 3: Snapshots of External Linked Data Projects & Activities | 28 |
| Appendix 4: Glossary of Terms | 32 |
| Appendix 5: Bibliography / List of Resources | 35 |

Executive Summary

This report documents the results and recommendations of the UC Libraries Linked Data Project Team's work. The Project Team's objective was to develop a deeper understanding of the use cases for, and potential benefits of, adopting a linked data approach to exposing and/or managing metadata, identifying infrastructure issues, and making recommendations for implementation opportunities at a systemwide level.

Linked data activities within the system are varied, ranging from awareness and study, to the development of large scale linked data systems and environments. This is not unexpected given the size, distribution, and diversity of the system. Perhaps more unexpected, is that cross-campus collaboration or coordination of linked data work is very limited within the system. The Project Team has identified a number of significant opportunities which would be made possible by systemwide implementation of linked data, leveraging both the UC Libraries' collective size as well as their individual diversity to better meet evolving user needs and develop innovative services.

As initial steps in realizing these opportunities, we recommend the following:

1. **Form a UC Linked Data Leadership Group** - As a foundational step we recommend the creation of a standing leadership group focused on developing the shared infrastructure and community of practice key to enabling the systemwide collaboration essential to realizing the opportunities identified in this report.
2. **Develop functional requirements for shared local authority infrastructure** – We also recommend the development of detailed functional requirements for a shared local authority infrastructure. This is an important area of work for meeting many of the practical challenges currently being encountered by UC Libraries, as well as being a topic of current interest in the larger library community.
3. **Consortial engagement with vendor services** – We recommend that mechanisms be established to ensure systemwide linked data requirements and needs are included in the California Digital Library's (CDL's) negotiations for related systemwide vendor services.
4. **Ensure Linked Data expertise in other related systemwide initiatives** – We also recommend that methods be established for coordinating systemwide linked data requirements and needs with other systemwide initiatives, such as the investigations into possibilities for a systemwide Integrated Library System (ILS) or Digital Asset Management System (DAMS).

These recommendations are practical initial steps focused on developing the essential shared infrastructure, common practices, and mechanisms for collaboration, needed for realizing many of the possible opportunities and benefits.

Background

The advent of linked data has provided libraries and other cultural memory stewards with opportunities to imagine improved access to the content within and beyond their collections - access that better affords navigation across the many relationships between resources that remain dormant when relying upon traditional cataloging approaches. Linked data approaches to managing and distributing metadata also hold the promise of improved efficiency in resource description and authority control workflows, for example, potentially allowing for new approaches to authority control that distribute application of expertise and avoid redundancy of effort. Since gaining currency among librarians over the past decade, many libraries have experimented with linked data approaches, and of those, some are embarking upon projects and programs that either operationalize linked data-informed workflows, deploy linked data enabled discovery environments, or both.

Prior to the work of the Linked Data Project Team, the UC Libraries have not as a system fully explored how and why linked data approaches, particularly those at a systemwide level, could benefit both our users and our staff. The Project Team asserts that linked data approaches to managing and/or exposing metadata can directly support the UC Libraries Strategic Priorities¹. Creating, maintaining, and/or sharing linked data directly supports the UC Libraries' vision of data-driven organizations leading development of innovative services, strategies, and systems. Linked data work seeks to fulfill the mission of enabling seamless, networked discovery and access, and managing the building blocks and products of scholarship and research in direct support of the University of California's teaching, learning, research, patient care, and public service goals. It also opens a range of opportunities for leadership and participation in national and global partnerships, and for development and optimization of shared services. Additionally, the advent of linked data technologies offers a unique opportunity to deepen collaborative relationships across the UC system. While collaborative relationships are not strictly necessary for employing linked data, they are essential for ensuring the reuse and sharing of data across traditionally disparate systems by supporting the development of interoperable policies and technological infrastructure.

In September 2017, the UC Libraries Direction & Oversight Committee (DOC) charged² a multi-campus Linked Data Project Team with developing "a deeper understanding within the UC Libraries of the potential benefits of adopting a linked data approach to exposing and/or managing metadata in various environments, and the infrastructure required to support implementation of linked data systemwide projects in the future." The Project Team's establishment and charge was directly informed by the November 2016 *UC Libraries DAMS Technology Report: Assessment of a Long-Term Solution for the UC*

¹ "Vision and Priorities." *UC Libraries*, <https://libraries.universityofcalifornia.edu/about/vision-and-priorities>. Accessed 29 Apr. 2018

² UC Libraries Direction & Oversight Committee. "Linked Data Project Team Charge." Sept. 2017. <https://libraries.universityofcalifornia.edu/groups/files/doc/docs/Linked%20Data%20Charge%20-%20Final%202017.pdf>.

*Libraries Systemwide DAMS*³ a key recommendation of which, accepted by both DOC and the UC Council of University Librarians (CoUL), was that a project team be formed to investigate linked data approaches. Concurrent to the work of the Linked Data Project Team, the Working Group for Systemwide ILS (SILS) Planning Project team was charged⁴ and established by CoUL, initiating work in June 2017 to investigate “how the UC Libraries might license a single, shared, systemwide ILS.” While these two teams have not had explicit interactions, at a minimum, findings from the Linked Data Project Team should be shared with SILS to the extent that such a system could support linked data workflows and/or discovery.

The Linked Data Project Team is comprised of librarians from across the UC Libraries with metadata and digital collections management expertise and with a range of experience with linked data approaches. Over the course of eight months, the team has:

- Compiled **use cases** for linked data approaches to both metadata management and exposure
- Conducted a **survey** of colleagues across the system to create and analyze a snapshot of UC Libraries linked data activities
- Conducted an **environmental scan** of significant current linked data projects external to the University of California
- Determined **opportunities** for the UC Libraries to explore shared infrastructure or access to shared services that capitalize on linked data approaches to address use cases and/or expand upon successful campus linked data activities
- Presented an initial analysis of UC Libraries linked data activities through a **birds of a feather presentation and dialogue at the UC Digital Library Forum (UCDLFx)** and following submission of this report, the team looks forward to sharing its findings and recommendations through a series of **webinars** for the UC Libraries community to convey potential benefits and applications of linked data

Linked Data Use Cases

1. Provide rich, interconnected resource description

Including links to external vocabularies and data sources in local data stores makes it possible to enhance the user experience by incorporating information from other sources as part of local

³ UC Libraries DAMS Technology Project Team. *UC Libraries DAMS Technology Report: Assessment of a Long-Term Solution for the UC Libraries Systemwide DAMS*. 18 Nov. 2016.
https://wiki.library.ucsf.edu/display/UCLDTP/UC+Libraries+DAMS+Technology+Project+Home?preview=/380146695/398659886/DOC_DAMS_FinalReport_18Nov2016.pdf#UCLibrariesDAMSTechnologyProjectHome-FinalReport.

⁴ University of California Council of University Librarians (CoUL), “Charge: Working Group for the Systemwide ILS Planning Project.” 21 Nov. 2017.
https://libraries.universityofcalifornia.edu/groups/files/coul/docs/CoUL_SystemwideILS_WorkingGroupCharge.pdf.

discovery activity. These enhancements can help users better interpret, understand, and contextualize library resources in a more interconnected discovery experience.

Example user stories:

- As a user I want to click on a link to Wikipedia so that I can learn more about “baile folklórico,” a term I found listed as a subject for a resource.
- As a user I want to quickly learn more about the “Smith, John” listed as an author by hovering over the name so that I can quickly tell if he is likely to be the biologist from the 1950s who studied mollusks that I am looking for without leaving the page for the item I am looking at.

2. Improved search and browse results and options

Leveraging the semantic relationships present in linked data can lead to numerous improvements in discovery within local systems by facilitating more robust browsing or searching for variants or related concepts. Of particular possible relevance to the UC system, given its multi-cultural composition and numerous international studies programs, is enabling multi-lingual and multi-script discovery. By incorporating information from external sources that compile multi-lingual and multi-script versions of entities and concepts, such as the Virtual International Authority File (VIAF) or Faceted Application of Subject Terminology (FAST), local systems can improve discovery by supporting searching for these entities and concepts by international communities.

Example user stories:

- As a user I want to search for Teddy Roosevelt and get results for resources which have “Roosevelt, Theodore, 1831-1878” as part of their record so that I get more complete results and do not need to repeat the search with multiple versions of the name.
- As a user I want to search for the term “sports” and be able to easily (in one to two clicks) include more specific kinds of sports like soccer, basketball, baseball, football, etc. in my search results so that I don’t have to manually search for “sports or baseball or soccer or basketball or football, etc.”
- As a user in East Asian Languages and Cultural Studies, I want to search for historic texts and their descriptions in Japanese so that I can conduct research directly in the language of my research.
- As a student writing a paper on political ballads, I want to explore related types of poetry that my library has available by browsing the broader category of ballads and then other subtypes of ballads so that I can discover other materials to help me develop my research topic.

3. Improve description and discovery of locally and regionally significant resources

Many entities described within local collections are not established in national authority files, and may lack the traditional literary warrant required to establish one according to the Name Authority Cooperative Program (NACO) guidelines. Often these entities have a regional or domain specific significance, and may be represented in collections spanning multiple UC campuses. These local authorities are traditionally managed in separate systems within each local environment, and are not typically shared between local systems or campuses. Making local linked data openly available for others to harvest and use can improve discovery and use of unique local content, both globally and within the UC system. Additionally, where existing Library of Congress names or subject headings do not reflect preferred terminology of diverse local communities, sharing local linked data authorities across the system is a potential means for standardizing use of the preferred terms.

Example user stories:

- As a processing archivist, I want to be able to share the information I gather about a local Chicana artist whose work appears in our archival and digital collections, but for whom there is not the literary warrant required for submitting a name heading to NACO so that colleagues across the system can use the information in their metadata work and improve discovery of resources by and about the artist.

- As a student in Native American studies, I want to search for library resources about Akimel O'odham communities using the communities' preferred name, "Akimel O'odham," rather than the Library of Congress Subject Heading (LCSH) "Pima Indians" so that I can find resources relevant to my research without having to search on a term I consider dated and offensive.

4. Improve discovery of UC collections across platforms and libraries

Incorporating linked data into otherwise disparate systems such as an ILS, DAMS, or archival system, libraries can create a more unified index of collections, and facilitate the discovery of resources regardless of the original system in which they are described.

By transforming local metadata into linked data and making it discoverable and consumable by outside systems such as search engines, local resources - especially rare or uniquely held resources - it could be much more easily discovered across a number of systems rather than siloed in local systems.

Example user stories:

- As a librarian responsible for collection development and considering systemwide collection-building opportunities, I want to understand how and where existing UC collections document Californians so that I can identify specific opportunities for collaboration.
- As a user interested in Medieval and Renaissance manuscripts, I want to be able to discover such manuscripts from a single point of search, whether they are described in a library catalog, as part of an archival collection, or in a digital repository, so that I can discover relevant resources with fewer queries in fewer locations.
- As a curator I want to be able to find materials relating to viticulture and enology from archives, library holdings, and digital collections across the system so that I can put together a multi-institutional exhibit on the history of vineyards and winemaking in California.

5. Publish UC Libraries' metadata for analysis and research

Just as large textual corpuses and data sets have been the focus of research using text and data mining methods, UC metadata, published as linked data, may also support analysis and research.

Example user stories:

- As a researcher I want to be able to access rich metadata for materials from archives and digital collections across the system so that I can do social network analysis of the early farm workers movement in California.
- As a humanities scholar I want to automatically access source metadata for science fiction collections to use in ongoing computational analysis and monitoring of the literature, so that I can identify emergent shifts and themes in the genre.

6. Improve discovery of library resources through search engines

Increased cross linking between library resources and other web content, as well as widened use of schema.org vocabularies on library web pages, could improve relevance ranking of library resources in search engines.

Example user story:

- As a user doing preliminary research on feminist epistemology in a web search engine, I want to discover more library resources in my results, so that I can find useful resources with fewer search queries in fewer locations.
- As a librarian I want to describe videos from our Holocaust Living History Workshop using the schema.org ontology so that they show up in Google as a rich snippet with the video thumbnails, which could lead to more searchers clicking on them.

7. Improve efficiency in metadata creation and management

Metadata creation and management work, and authority management work in particular, in libraries is often siloed requiring work to be duplicated in separate systems. A distributed authority management model offers a number of potential efficiencies, including automation of maintenance tasks and ability to reuse local authorities across multiple, otherwise disparate library systems like ILS, DAMS, and archival systems. Such a model could also make it possible for individuals across UC Libraries to leverage their subject expertise to create and contribute to shared authority records, rather than limiting participation in authority work to primarily catalogers with specialized NACO and Subject Authority Cooperative Program (SACO) training.

Example user stories:

- As a special collections cataloger I want to be able to access and use the same headings for rare books as those used for related archival descriptions so that I do not need to create a new record and so the records in both the archival management system and ILS use the same heading.
- As a processing archivist I want to be able to take better advantage of information gathered by our visual resources cataloging specialist about an emeritus visual arts professor whose heirs recently donated his papers to the library so that I can reduce my research and processing time.
- As a SACO contributor I want to be able to use information gathered by cataloging staff with domain expertise (who are not SACO trained) so that I can more easily put together a subject heading proposal for submission.
- As a cataloger I want changes made by the Library of Congress to be reflected in our catalog automatically so that I do not have to manually update them in the database, or pay an authority vendor to make the changes for me.

University of California Libraries Linked Data Activity

Linked data activities within the UC Libraries reflect the variety of approaches towards linked data investigation, experimentation, and implementation being taken within the wider library community. To better understand the existing state of linked data activity across UC Libraries, this group surveyed colleagues ([Appendix 2](#)) across the system regarding past, current and planned linked data activities; domains, systems and tools involved; and challenges faced. The following is a summary of the various linked data activities undertaken within the UC Libraries, including note of the use case(s) they support. A list of projects is included in [Appendix 1](#).

In linked data projects originating in the bibliographic domain, activities have often focused on large-scale collaboration and experimentation, and involved partnering with vendors and/or external organizations. Berkeley is participating with 16 other institutions in the SHARE Virtual Discovery Environment (SHARE-VDE) project led by Casalini Libri, involving the mapping of several million MARC bibliographic and authority records to Resource Description Framework (RDF) for use in a prototype linked data discovery system. In a similar vein, Davis has worked with a number of vendors toward the creation and implementation of a fully linked data-enabled catalog, including participation in BIBFLOW and OCLC's Linked Data Wikibase Prototype pilot. The large scale and holistic scope of these projects could possibly support any and all of our identified use cases.

As part of the PCC, Los Angeles and Davis, along with 11 other institutions, are part of a pilot project investigating International Standard Name Identifiers (ISNI) and determining if ISNI could be incorporated into libraries' authority work. By creating ISNIs for local identifiers, augmenting local identifiers with ISNI information, and reconciling local identifiers against a global database, this project could support more rich, interconnected resource descriptions, especially for locally and regionally significant resources (1 & 3), improved searching and browsing in local systems (2), and improved efficiencies in metadata creation and management (7).

Los Angeles and Davis participated in a task force focused on the mapping of serials MARC metadata into the Bibliographic Framework Initiative (BIBFRAME) ontology overseen by the PCC, and San Diego and Los Angeles are working with several institutions on an MLA project to develop a BIBFRAME based ontology for performed music. Ensuring that MARC metadata is accurately mapped to linked data ontologies, identifying areas in need of improvement, and establishing a mapping that could be widely implemented, supports improved discovery of UC collections across platforms, including search engines (4 & 6).

Yet activities regarding Machine-Readable Cataloging (MARC) metadata are not limited to these larger-scale or planning-focused efforts. For example, Santa Cruz is working with Backstage Library Works to incorporate authority Uniform Resource Identifiers (URIs) into subfield \$0 for various fields in MARC records. This is intended to make MARC metadata more "linked data ready," and richer and more interconnected (1). Irvine's Artists' Books Project piloted the transformation of MARC metadata to a linked data format, and created a prototype visualization tool (2) to improve discovery.

Linked data work originating in the digital repository domain has often focused on implementing smaller-scale local projects in existing systems, sans vendor involvement. The activity undertaken by the greatest number of campuses was reconciliation of entities and terms in digital repositories with data or authorities from external systems, with Riverside, San Diego, Santa Barbara, Santa Cruz, and CDL all reporting work in this area. These projects employ a variety of tools, such as OpenRefine, Python, SPARQL Protocol and RDF Query Language (SPARQL), and spreadsheets, to connect terms in local systems to vocabularies and data sources such as the Library of Congress Name Authority File (LCNAF), LCSH, FAST, VIAF, and Wikidata in order to create more interconnected authorities (1) for local resources (3) and possibly establish a more robust and distributed workflow for creating and maintaining authorities (7). San Diego also cited providing links to Wikipedia (1) as a goal. Another activity involving digital repositories was work undertaken by Merced to create a mapping between the Solr index in Calisphere and schema.org. This work is intended to improve the discoverability and display of Calisphere collections and items in search engines (6).

There are also a number of current CDL persistent identifier services relevant to linked data work, including the Archival Resource Key (ARK) identifier standard, the Names2Things (N2T) ARK resolver service, and the EZID identifier creation and management service, all of which are intended to support a more efficient way of creating and managing authorities (7) in an interconnected, global environment (1 & 3).

Finally, the majority of UC Libraries are involved with learning more about linked data. Some campuses have organized study or discussion groups on linked data in preparation for future linked data projects or implementations.

While UC Libraries have implemented a number of linked data activities, some challenges remain. Two common issues identified by the survey were getting institutional support for undertaking linked data activities, and finding ways to prioritize linked data activities when they may directly compete for time with other essential job duties.

External Linked Data Activity

A central component of the bibliographic metadata communities' move towards linked data is BIBFRAME, which has made substantial progress in developing shared data models and schemas to support the description of bibliographic resources. As noted in UC campus activities, this includes the extension of BIBFRAME ontologies to better support specific subject domains and resource types. It also involves developing methods for large scale conversion of extant data. A few examples include the Library of Congress' MARC to BIBFRAME Conversion tool and Stanford's Linked Data for Production (LD4P) Tracer Bullets and Data-Pipeline projects. There are also a number of vendors developing conversion and reconciliation services. Of potential particular interest for the UC System are OCLC's Linked Data Wikibase Prototype and the SHARE-VDE project, both of which include UC campus involvement in their testing, and which are working with large data sets.

There are also a number of linked data based digital repository systems being developed. The transition of the widely adopted Fedora repository to a W3C linked data repository platform is a key development in this area. It is a key layer which the Samvera and Islandora repository applications are frequently deployed on. The Fedora, Samvera, and Islandora communities have been working in a number of key linked data areas, ontologies and vocabulary alignment, modeling of complex RDF modeling, and development of community and co-development models. Straddling both digital repository and bibliographic data stores, the VIVO and Vitro applications provide a number of ontology editing, metadata creation, and repository management capabilities integrated in an end-user application interface. The Questioning Authority Gem, developed as part of the Linked Data for Libraries (LD4L) project, provides integrated term look up for cataloging and description. Designed with the goal of being extensible to multiple repositories (VitroLib and Samvera currently), it integrates look up functionality of terms from external vocabularies into metadata creation user interfaces, as well as brokering information to and from multiple vocabularies with varied encoding and querying standards.

Many major national or multinational libraries and knowledge organizations have published linked data sets, including the Library of Congress, the Swedish Union Catalog, the British National Bibliography, the Getty, and Europeana. Their work provides sources for authoritative LD vocabularies and data, but also models which can inform development of technical requirements for implementing and maintaining a robust, large scale, linked data publication infrastructure.

There is also a great deal of interest within the library community in developing new models for cooperative and shared authority creation and management, as illustrated by the recent National Strategy for Shareable Local Name Authorities National Forum.⁵ PCC has been actively engaging in exploring alternative ways to work with authorities and identifiers from multiple systems, such as LCNAF, ISNI, VIAF, Open Researcher and Contributor ID (ORCID), and local systems, to discover ways of managing identifiers in a collaborative environment, and determining how current authority practices may change in a more identifier-focused environment. Authorities form a productive point of intersection between bibliographic, archival, and digital library domains, and a robust means of connecting to non-library tools and resources.

While network graph based access tools like UC Irvine's Artists' Books Discovery Tool or the Big Data Infrastructure Visualization Application (BigDIVA) interface are the most unique and visually recognizable User Interface (UI) application of linked data, many of the most widespread uses are less obvious, utilizing the behind the scenes data linkages to improve the user experience without fundamentally changing it. For example, Google's Knowledge Graph cards present a summary of information about an entity to a user, as well as alternative possible entities with the same string name. While these summaries could be manually created for a very small percentage of entities, the scale needed is possible because of linked data. While specifics about the technology behind Google's Knowledge Graph are not public, they introduced it in 2012 as a graph "that understands real-world entities and their relationships to one another: things, not strings," containing "more than 500 million objects, as well as more than 3.5 billion facts about and relationships between these different objects."

⁵ Casalini, Michele, et al. *National Strategy for Shareable Local Name Authorities National Forum : White Paper*. report, 29 Mar. 2018, <http://ecommons.cornell.edu/handle/1813/56343>.

The British Broadcasting Corporation (BBC) also uses information harvested (e.g. biography, discography, YouTube channel link, etc.) from linked data sources, in combination with data from BBC media sources, to present users with an information-rich experience which would not be possible if only BBC created data was used.

The larger linked data landscape shows a diversity of approaches, but with notable shared areas of concern in the areas of data conversion and reconciliation, development of platforms, evolving approaches to authorities creation and management, and utilization of linked data for improving and enhancing the user experience.

Systemwide Opportunities

Opportunity 1: Shared local authority management

As libraries across the UC system encounter common challenges around authority management in a linked data environment, there is simultaneously a growing national conversation around leveraging and sharing local authorities as linked data. At the system level, UC is well-positioned to engage in such efforts that can both address immediate pragmatic challenges in campus linked data workflows, and contribute meaningfully to a developing area in the larger linked data landscape.

Because linked data by definition uses URIs to represent concepts and terms, management of and access to these identifiers is a central dependency of nearly all use cases outlined in this report. Existing tools and workflows tend to lean heavily on centralized controlled vocabularies and authority management infrastructure maintained by the Library of Congress, the Getty, and OCLC, without straightforward solutions for creating, managing and reconciling identifiers for local concepts and terms that fall outside the scope of such national and international systems. At the campus level, the challenge of managing local authority data as linked data can be a stumbling block that prevents scaling-up of linked data efforts, due to the web infrastructure it requires and complex data architecture questions it can surface.

Three recent Institute of Museum and Library Services (IMLS) funded exploratory projects have addressed interrelated facets of shared local authority management: at a national scale, the National Strategy for Shareable Local Name Authorities National Forum brought stakeholders together to develop a minimum viable specification for local identity data management⁶, while the Western Name Authority File (Mountain West Digital Library) has begun to tackle linked data authority management at the regional digital collections level⁷; and Florida State University's "*Towards engaging researchers in research identity data curation*" project addressed the design of scalable, reliable infrastructure for

⁶ Michele Casalini et al., "National Strategy for Shareable Local Name Authorities National Forum : White Paper," report, March 29, 2018, <http://ecommons.cornell.edu/handle/1813/56343>.

⁷ University of Utah. "Linking People: Developing Collaborative Regional Vocabularies." *Institute of Museum and Library Services, Award #LG-72-16-0002-16*, 2016, <https://www.ims.gov/grants/awarded/lg-72-16-0002-16>.

researcher identity management (such as ORCID) and important questions around user buy-in.⁸ These efforts lay important groundwork and point to much needed work ahead.

Anticipated benefit

Shared local authority management would grow linked data capacity and infrastructure across the UC system, with potential to create both a rich linked data resource in itself and facilitating a wide array of the use cases outlined in this report, such as enhancing resource discovery with information from related sources, improving metadata creation and management workflows, and publishing local data to support external reuse and discovery. Its uses span both internal and external audiences, across bibliographic, archival, and digital collections.

Critical factors

Data modeling: Successfully sharing local authority data at the system level requires establishing a shared data model for maximal interoperability of authority data from multiple sources.

Infrastructure: Successfully sharing local authority data at the system level requires robust, scalable technical infrastructure to support identifier persistence, query, and disambiguation services for a large volume of data frequently updated from distributed sources.

Governance and policy: Successfully sharing local authority data at the system level requires establishing a governance model that involves stakeholders across the system in making decisions around critical issues such as establishing the aforementioned data model for authority data, funding models and financial sustainability, data licensing, identity data privacy and confidentiality, and responding effectively to community needs throughout the life of the project or service.

Opportunity 2: Consortial engagement with vendor services

With a number of library vendors now offering or exploring linked data services, and several UC libraries having engaged vendors for linked data efforts both small- and large-scale, the system has the opportunity to take a collaborative approach to evaluating and engaging such services at the consortial, system level.

Because the linked data vendor sector is relatively young, “out-of-the-box” products and services are few, and library use of vendor services in this space has frequently been structured as pilot projects or exploratory partnerships. Particularly by engaging at a system level, this may be an opportunity to help shape the future of vendor-provided products and services and advocate for library values, goals and interests to form more equitable and beneficial partnerships with vendors, including those that offer targeted linked data-related services as well as those offering suites of services that are just beginning to, or could potentially, address linked data ingest, management, and/or publication.

⁸ Stvilia, Besiki, et al. “Toward Engaging Researchers in Research Identity Data Curation.” *Institute of Museum and Library Services, Award #LG-73-16-0006-16*, 2016, <https://www.ims.gov/grants/awarded/lg-73-16-0006-16>.

A consortial approach could expand linked data opportunities, particularly for smaller campuses, and would bring a more varied set of contexts, collections and concerns to bear on evaluating (and advocating for) vendor offerings than engaging at the single-campus level.

Anticipated benefit

A strategic system-level approach to engagement with vendor services increases the likelihood that the use of such services will support UC Libraries goals, values, and policies and be cost-effective and beneficial to libraries across the system in the long term. This approach also invites a higher level of alignment and collaboration between campuses in linked data work.

Critical factors

Data stewardship: Successfully engaging vendor services at the system level requires careful consideration of issues surrounding data stewardship and autonomy, ensuring that UC Libraries continue to be able to make metadata available per the spirit of openness that is expressed in the UC Libraries Metadata Sharing Policy⁹. Any vendor relationship must support sharing metadata as widely as possible as outlined in the UC Libraries Metadata Sharing Policy, and in support of Goal 2 of the UC Libraries Vision and Priorities (“Maximize discovery of and access to information resources”).

Cost: Successfully engaging vendor services at the system level requires careful consideration of associated costs, strategically targeting areas where engaging vendor services can be most effective.

Governance and policy: Successfully engaging vendor services at the system level requires establishing a governance model that involves stakeholders across the system in making decisions around critical issues such as goals, requirements that reflect agreed upon best practices, budget, and evaluation.

Opportunity 3: Collaborative development

The collaborative development model recommended by the UC Libraries DAMS Technology Report has potential to be highly beneficial in a linked data environment by leveraging expertise and resources across the system to more quickly meet emerging needs and develop innovative services.

The potential of this benefit may be most significant in the user-facing systems, where the larger library community is in a relatively early stage of development. No matter how robust or well-maintained the underlying data, realizing many of the user-facing promises of linked data will not be possible without investing in the development of technical infrastructure to support discovery across repositories as well as User Experience (UX) and UI design for leveraging linked data capabilities.

Given the diversity and richness of systemwide resources (both human and collections), UC Libraries may be in a unique position to lead design and development of linked data and semantic web discovery

⁹ Metadata Policy Project Team. “University of California Libraries Metadata Sharing Policy.” *UC Libraries*, 2015, https://libraries.universityofcalifornia.edu/groups/files/sag2/docs/Final_UC_Libraries_Metadata_Sharing_Policy_May8_2015.pdf.

applications. Collaborative development would greatly increase the impact for development efforts, benefitting all campuses and users.

Such an approach might be adopted in conjunction with adoption of shared systems or technology stacks, such as a systemwide ILS or systemwide DAMS. Shared infrastructure could simplify collaborative development of reusable tools, enhancements, and processes across the system. However, centralized data aggregation is not a prerequisite for a shared development approach, particularly in working with interoperable linked data models.

Anticipated benefit

Evaluating, and then acting upon, opportunities to engage in shared development as an alternative or complement to vendor provided systems and services will leverage linked data, for example, to improve user discovery and access to information across all campuses. Engaging in shared development efforts to realize the potential of linked data may offer a particularly high-impact role for UC Libraries in a less-developed and much-needed area of library technology.

Critical factors

Resources and funding: As stated in the UC Libraries DAMS Technology Report, successfully adopting a shared development model “will require a formal, collective commitment to the long-term resources required for both development and ongoing operations.”

Governance and structure: As stated in the UC Libraries DAMS Technology Report, successfully adopting a shared development model will require defining a structure and governance model that “achieve[s] desirable economies of scale, while still providing flexibility and local autonomy to meet campus goals.” Opportunities to leverage linked data through shared development efforts provide yet another example of the need for such a structure and governance model

Opportunity 4: National and international collaboration

Linked data poses both pragmatic and conceptual questions that libraries around the world are working to answer, often from a perspective of experimentation, development, and emergent best practices. Strong collaborative relationships are both a vehicle for capacity-building and a natural fit for an area of practice that values openness, interoperability, and expanded discovery.

UC participation in large-scale linked data initiatives has been present, but rarely a prominent focus. Opportunities for growth are present both in terms of supporting and strengthening systemwide and multi-campus efforts, as well as strengthening UC’s presence as a system in national and international linked data initiatives.

Many prominent institutional voices in the U.S. library linked data landscape are private universities. As a diverse public university system, UC can likely bring insights and perspectives to these spaces that are currently underrepresented, benefiting both our own work and advancing the goals for and approaches to linked data implementation as a whole.

Anticipated benefit

Strengthening UC participation in external collaboration will improve the efficacy and impact of UC Libraries linked data work and contribute to the large-scale advancement of the goals for linked data in libraries.

Critical factors

Intra-system collaboration and communication: Successfully strengthening UC participation in external collaboration will require strengthening collaboration and communication within and across the system. Regularly maintained and updated information on linked data projects throughout the system would enable UC library administrators and staff to easily stay up to date on systemwide work, represent that work to external colleagues, and identify new opportunities for collaboration.

Planning: Successfully strengthening UC participation in external collaboration will require thoughtful planning and systemwide coordination of effort in order to grow beyond the existing ad hoc approach.

Support: Successfully strengthening UC participation in external collaboration will require supporting such efforts as core to UC Libraries mission, rather than matters of individual or single-campus interest.

Recommendations

The University of California Libraries are well-positioned to respond at a system level to a variety of linked data opportunities that could positively impact our libraries and users in the following ways:

1. Improve user experience.
2. Streamline technical services workflows.
3. Improve the quality of library metadata.
4. Improve interoperability of cataloging, archival, and digital library systems; thereby improve management and access to resources.
5. Improve our ability to respond effectively to global changes in information ecosystems.

Because large-scale linked data work requires significant investment of labor and technical infrastructure, the following recommendations are intended as pragmatic initial steps which do not address the full range of opportunities identified in the previous section.

To this end, this report outlines one (1) foundational and three (3) specific recommendations. Recommendation 1, the foundational recommendation, speaks to a need to foster productive collaboration across the UC system, and serves as necessary groundwork for the remaining Specific Recommendations.

Recommendation 1: Formation of a standing UC Libraries Linked Data Leadership Group

We recommend that CoUL/DOC form a standing UC Linked Data Leadership Group with the charge to define and coordinate technical specifications, functional requirements, approaches and practices to guide linked data engagement across the system.

The rationale behind the formation of such a group is to provide a stable and ongoing mechanism for supporting linked data collaboration in the context of a large, complex system comprised of geographically dispersed partners with distinct local contexts, needs and priorities.

The UC Linked Data Leadership Group would provide an official structure for system partners to define common linked data goals, and the best practices and functional requirements that will help successfully meet those goals. With these in hand, campuses can work independently on linked data applications and workflows while positioning this work to function in the systemwide linked data ecosystem (and beyond) as it evolves.

In addition to defining a common framework for Linked Data development and implementation The UC Linked Data Leadership Group would also serve as a needed consultancy point for other working groups (such as, for example, the Systemwide Integrated Library System group) to consult on aspects of their work that either impact or are impacted by Linked Data.

Important economies of scale and efficiency can also be achieved by direct co-development and collaboration in this area. As such, a primary purpose of the the UC Linked Data Leadership Group will also be to support productive and defined structures for direct collaboration. To this end, the group in consultation with CoUL and/or DOC, can form and fill membership in working groups, focused on specific tasks. These could include, for example, exploratory working groups, cooperative development initiatives, and groups focused on defining functional specifications.

Group membership should ideally include an equal number of participants from each campus library and the California Digital Library. The group should include archival, digital, and bibliographic expertise, and both metadata and technical specialists.

It is recommended that the UC Linked Data Leadership Group scope include:

1. Creation of annual inventory of UC Library linked data activities and major extra-UC collaborative efforts
2. Identification and prioritization of community beneficial opportunities for development or purchase of tools or services
3. Recommendations for community linked data best practices, such as:
 - Minting and reconciling URIs
 - Inclusion of URIs in existing record and resource management systems
 - Shared authority record data model

4. Creation of linked data related functional specifications and requirements to be used for:
 - Soliciting, evaluating, and negotiating contracts for vendor services
 - Planning and implementation for new development projects
 - Enabling better cost and feature comparison between vendor services and development

Specific goals and deliverables will depend on CoUL/DOC identified systemwide prioritization, and the related available level of support for the members expected time commitments and funding for in-person meetings. While much of the work of the UC Linked Data Leadership Group will be accomplished virtually, regular, in-person meetings are important to developing strong collaborative working relationships and maintaining progress. Although not required for the formation of standing UC Linked Data Leadership Group, financial support at the system level would enable broader participation across the system than would depending solely on individual campus', and so could be an important consideration for achieving wider engagement, participation, and implementation. Acknowledging that this system level of financial support may not be currently feasible, it is recommended for CoUL/DOC's consideration in ongoing evaluation and prioritization of goals.

Recommendation 2: Development of functional requirements for shared local authority infrastructure

Successful systemwide collaboration in the area of local authority management, as identified in Opportunity 1, depends upon the adoption of a core set of common practices. For example, what are the minimum data elements that must be associated with an entity to be considered an authority? How do we approach reconciliation across vocabularies and domains? What are the common technical requirements and practices for creating, maintaining, and managing URIs as identifiers for authority entities?

We recommend that the system perform a deep-dive investigation of available models of implementing systems to support a common, collaborative infrastructure for supporting linked data including URI creation and management, reconciliation tools and services, and a shared authority file. In the event that Recommendation One as described above is adopted, this investigation should be conducted by a sub-working group under the authority of the UC Linked Data Leadership Group. Alternatively, a standalone working group authorized by CoUL/DOC or an authorized CKG could perform this function. In either case, this group should be charged with developing functional specifications and producing cost and feasibility analysis of various potential solutions, ranging from co-development to adoption of vendor-provided services.

Recommendation 3: Include systemwide linked data needs and requirements in consortial engagement with vendor services

As outlined in Opportunity 2, consortial engagement with vendors provides an opportunity to advance UC Libraries goals and priorities in regard to linked data. Because CDL is already well-established as a site of consortial engagement with external vendors, it is our recommendation that CDL continue to serve in this capacity. CDL should formally consult with the UC Linked Data Leadership Group in the development

and review of functional requirements for metadata-related products and services acquired by CDL on behalf of the system. In the event that Recommendation One above is not accepted by CoUL/DOC, CDL should be directed to formulate a specific plan for engaging with individual campuses in this regard.

The working group will leverage their expertise in seeking to “future-proof” vendor services engagement, ensuring that contracting decisions will not lock UC Libraries into systems that preclude linked data management and functionality that the system determines to be in our and our users’ best interest. This structure will also allow individual campus libraries to have direct input in decision-making around critical issues, including requirements and strategic goals, while allowing CDL to engage vendors autonomously.

Recommendation 4: Ensure Linked Data expertise in membership of relevant systemwide initiatives

We recommend that CoUL/DOC takes steps to ensure that any groups whose purview includes practices or technology related to metadata creation and/or management directly and formally engage with the UC Linked Data Leadership Group as part of their activity in order to ensure that the specific recommendations of those groups align with systemwide best practices in the area of linked data. In the event that a UC Linked Data Leadership Group not be formed, we recommend that CoUL/DOC insure that the membership of all relevant groups include identified Linked Data expertise.

At present, there are two such groups currently active:

- A. Working Group for the Systemwide ILS Planning Project:** Given the time, effort, and resources involved in selecting and implementing a system at this scale, it is important to evaluate the candidate SILS systems for compatibility with the linked data needs and goals of the system. The SILS project should be directed to ensure that the relevant SILS Expert Groups to be formed during this phase of work (for example, groups focused on cataloging, metadata, and discovery) include members with identified Linked Data Expertise. .

- B. Fedora DAMS Working Group:** CoUL/DOC has currently charged a working group to investigate the adoption of a systemwide Fedora DAMS repository. Linked data efforts should actively advance the agenda of connecting our heretofore disconnected digital asset, archival, and traditional catalog data stores. To this end, we recommend that the Shared Fedora Development Working Group be directed to review both this report and the publications of the Linked Data Best Practices Working Group as recommended above, or other groups that may be formed, as guidelines to ensure that their work reflects consortially adopted best linked data practices. It is also recommended that membership in the Shared Fedora Development Working Group be expanded to include two representatives from the Linked Data Best Practices Working Group or individual members with identified Linked Data expertise

Appendix 1: Snapshot of UC Libraries Linked Data Projects & Activities

Authority control, reconciliation, identity management

Consuming Big Linked Open Data in Practice: Authority Shifts and Identifier Drift *

An exploratory project in which bulk data from VIAF was evaluated for integration into HathiTrust metadata to improve discovery and collection management.

Documentation:

<https://code4lib.org/2014/12/05/consuming-big-linked-open-data-in-practice-authority-shifts-and-identifier-drift.html>

Campus: CDL

Contact: Kathryn Stine

Status: Completed

Social Networks and Archival Context (SNAC)

Project focused on extracting descriptions of names (persons, families and corporations) from archival records, disambiguating identities, and linking to related resources and entities. SNAC, in part, enables archivists, librarians, and scholars to jointly maintain information about the people documented in archival collections.

Documentation: <http://snaccooperative.org/static/about/about.html>

Campus: CDL (see additional SNAC collaborators here:

<http://snaccooperative.org/static/about/collaborators.html>)

UC Contact: Rachel Hu (member of the SNAC Standards and Editorial Working Group and Communications Working Group)

Status: Ongoing

Reconciliation and authority creation for Digital Collections *

Projects focus on authority management and metadata enrichment for all local authorities in the Digital Asset Management System. Work includes conversion of LCSH pre-coordinated subjects to “faceted” or atomic terms, reconciliation, deduplication, and enhancement with links to external data stores (i.e., Wikidata and Wikipedia). Several campuses are engaged in similar projects at different scales.

Campuses: San Diego, Santa Barbara, Santa Cruz & Riverside

Contacts: Arwen Hutt (San Diego), Chrissy Rissmeyer, Tom Brittnacher & David Seubert (Santa Barbara), Rachel Jaffe (Santa Cruz), Noah Geraci (Riverside)

Status: Underway

PCC ISNI Pilot *

Developing an understanding of ISNI tools and systems, creating PCC documentation and training for its members, and putting in place member supports for cost-effective scaling of broader PCC involvement in ISNI.

Documentation: <https://wiki.duraspace.org/display/PCCISNI/PCC+ISNI+Pilot+Home>

Campus: Los Angeles & Davis

Contact: Paul Priebe (Los Angeles) & Xiaoli Li (Davis)

Status: Underway

EZID

EZID is a CDL service that offers a shared infrastructure for minting and managing URIs for research outputs. Includes API for automating use. A number of open source applications have been developed for connecting EZID services to different systems and architectures, including Samvera, DSpace and Crossref.

Documentation: <https://ezid.cdlib.org/learn/>

https://ezid.cdlib.org/learn/open_source

Campus: CDL

Contact: Joan Starr

Status: Completed

ARKs and ARKs in the Open Project

The Archival Resource Keys (ARKs) standard is a persistent identifier standard developed at CDL. ARKs in the Open Project is a collaborative project recently started with DuraSpace to develop a wider community of support and long-term sustainability for the standard and infrastructure.

Documentation: <https://wiki.duraspace.org/display/DSP/ARKs+in+the+Open+Project>

Campus: CDL

Contact: John Kunze

Status: Completed (ARKs), In progress (ARKs in the Open Project)

N2T (Name-to-Thing) Resolver

Persistent identifier resolver which stores redirection and descriptive information for individual identifiers.

Documentation: <http://n2t.net/e/about.html>

Campus: CDL

Contact: John Kunze

Status: Completed

Development and/or implementation of linked data based systems

Implementation of Fully Linked Data Enabled Catalog *

Includes the ability to publish and ingest data in MARC with URI, BIBFRAME, and schema.bib formats. Project team is collaborating with vendors OCLC and ExLibris.

Campus: Davis

Contact: Carl Stahmer

Status: Underway

DAMS 4 / Digital Collections

Development and implementation of a linked data based digital repository system (UCSD DAMS v4). Consisting of a Samvera based access layer, triplestore, using a locally developed RDF data model, ARK identified local authority records, and a separate ingest and management application. Plans underway for DAMS v5.

Documentation:

<https://library.ucsd.edu/research-and-collections/collections/digital-collections/dams.html>

Campus: San Diego

Contact: Declan Fleming

Status: Completed

Alexandria Digital Research Library (ADRL) *

Development and implementation of a linked data based digital repository system. Within the system, structural metadata is expressed using the Portland Common Data Model (PCDM) and descriptive metadata is expressed using a local RDF based model.

Documentation:

- <https://drive.google.com/drive/folders/0BxNsbR6fxrHHcnJ3QzRSM2RBS2c>
- https://docs.google.com/spreadsheets/d/1D4Kd92HpWiwxQssEA15bWsMIO_Zguw3lgDhITCfZ9ik/edit
- <https://pcdm.org/2016/04/18/models>

Campus: Santa Barbara

Contact: Chrissy Rissmeyer & John Ajao

Status: Underway

Linked Data Digital Repository System *

Development and implementation of a linked data based digital repository system. Within the system, structural metadata is expressed using the Portland Common Data Model (PCDM) and descriptive metadata is expressed using a local RDF based model.

Campus: Santa Cruz

Contact: Sue Perry

Status: Underway

Collaborative data modeling for co-development

Exploratory proof-of-concept project to determine the minimum level and type of alignment required between Samvera/Fedora4/Portland Common Data Model (PCDM) based data models, to enable co-development or code repurposing. Includes: detailed analysis of UCSB and UCSD data models, exploration of existing tools and processes to support the collaborative development and maintenance of such a hybrid data model, and identification of potential governance requirements.

Campuses: Santa Barbara & San Diego

Contact: Chrissy Rissmeyer (Santa Barbara) & Arwen Hutt (San Diego)

Status: Underway

Conversion of bibliographic data to linked data

SHARE-VDE *

UC Berkeley is a lead in the SHARE-VDE project, working in partnership with Casalini Libri and other institutions. SHARE-VDE is an effort to engage in large scale (+40 million bibliographic records) conversion, aggregation, publication and discovery of bibliographic linked data.

Documentation: <http://share-vde.org>

Campus: Berkeley

UC Contact: Erik Mitchell

Status: Underway

Linked Data for Performed Music *

Participation in the Music Library Association's Linked Data Working Group. Working in partnership with the Library of Congress, the Program for Cooperative Cataloging and other institutions, this group is developing a BIBFRAME based ontology for performed music. This project is part of the Mellon Foundation funded grant Linked Data for Production.

Documentation: http://www.musiclibraryassoc.org/mpage/cmc_ludwig

Campuses: Los Angeles & San Diego

Contact: Hermine Vermeij (Los Angeles) & Jim Soe Nyun (San Diego)

Status: Underway

Mapping the CONSER Standard Record onto BIBFRAME *

Working in partnership with the Library of Congress, the Program for Cooperative Cataloging and other institutions, this project sought to identify how well BIBFRAME can accommodate CSR elements, and make recommendations regarding areas in need of improvement or further investigation.

Documentation:

<http://www.loc.gov/aba/pcc/bibframe/TaskGroups/CSR-PDF/FinalReportCONSERtoPCCBIBFRAMETaskGroup.pdf>

Campuses: Los Angeles & Davis

UC Contact: Kevin Balster (Los Angeles) & Xiaoli Li (Davis)

Status: Completed

Mapping bibliographic (serials) metadata from MARC21 to BIBFRAME *

The primary goal was to assess the robustness of the BIBFRAME schema when used in describing serials.

Campus: Los Angeles

Contact: Kevin Balster

Status: On hold

Enhancement of Bibliographic Metadata *

The library contracted with Backstage Library Works to provide URIs in subfield 0 for 1xx, 6xx, and 7xx fields in bibliographic records and 1xx fields in authority records.

Campus: Santa Cruz

Contact: Marcia Barrett

Status: Completed

Discovery & data visualization

Piloting Linked Open Data for Artists' Books; Artists' Books Discovery Tool

This National Endowment for the Humanities funded project used linked data to create a visualization representing artists' books in UCI Libraries' Special Collections & Archives collection.

Documentation: <https://www.lib.uci.edu/sites/all/plodab/about.php>

Campus: Irvine

Contact: Laura Smart

Status: Completed

Schema.org markup for Calisphere *

Implementing schema.org in Calisphere pages in order to improve discovery and display of objects in search engine results.

Documentation:

<https://wiki.library.ucsf.edu/display/UCLDC/Calisphere+schema.org+implementation>

Campuses: Merced & CDL

Contact: Emily Lin (Merced)

Status: Completed

Exploration & experimentation

Survey respondents at several campuses (Berkeley, Santa Barbara, CDL, Los Angeles & San Francisco) reported that they have or are currently engaged in self or group study on topics related to linked data including BIBFRAME and RDF. Additionally, Rhonda Super (Los Angeles), created a Semantic Web LibGuide (<http://guides.library.ucla.edu/semantic-web>) which pulls together training materials, best practices, use cases, vocabularies, tools, etc.

These snapshots focused primarily on activities of UC Libraries. Linked Data activities may also well be underway across the system outside of the libraries, such as work undertaken by UCSF on the Open Knowledge Network initiative: <http://ichs.ucsf.edu/open-knowledge-network/>.

Appendix 2: UC Libraries Linked Data Survey

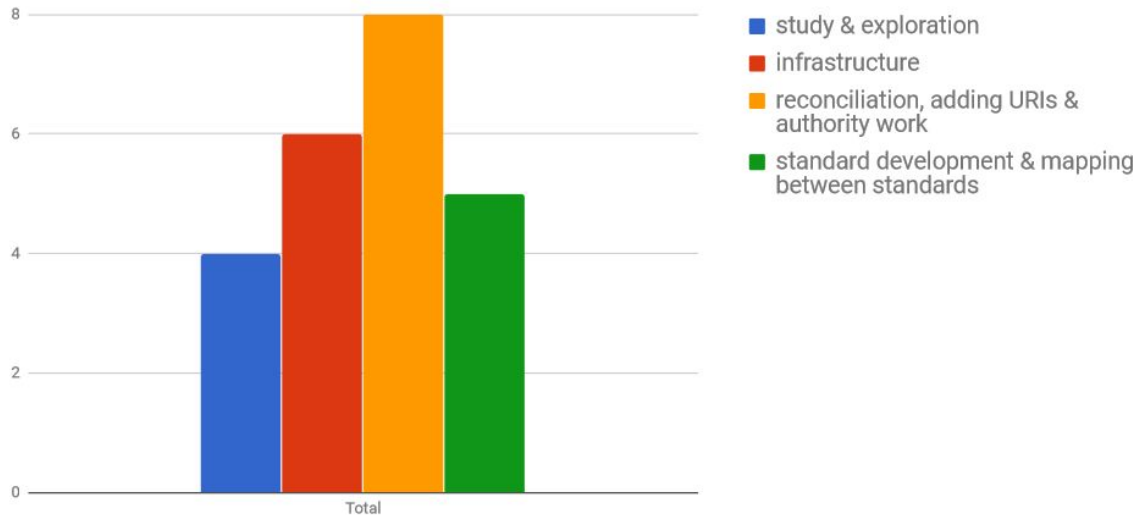
In the winter of 2018, the UC Linked Data Project Team conducted the following survey to identify and gather information on linked data activities in the UC system.

UC Libraries Linked Data Survey

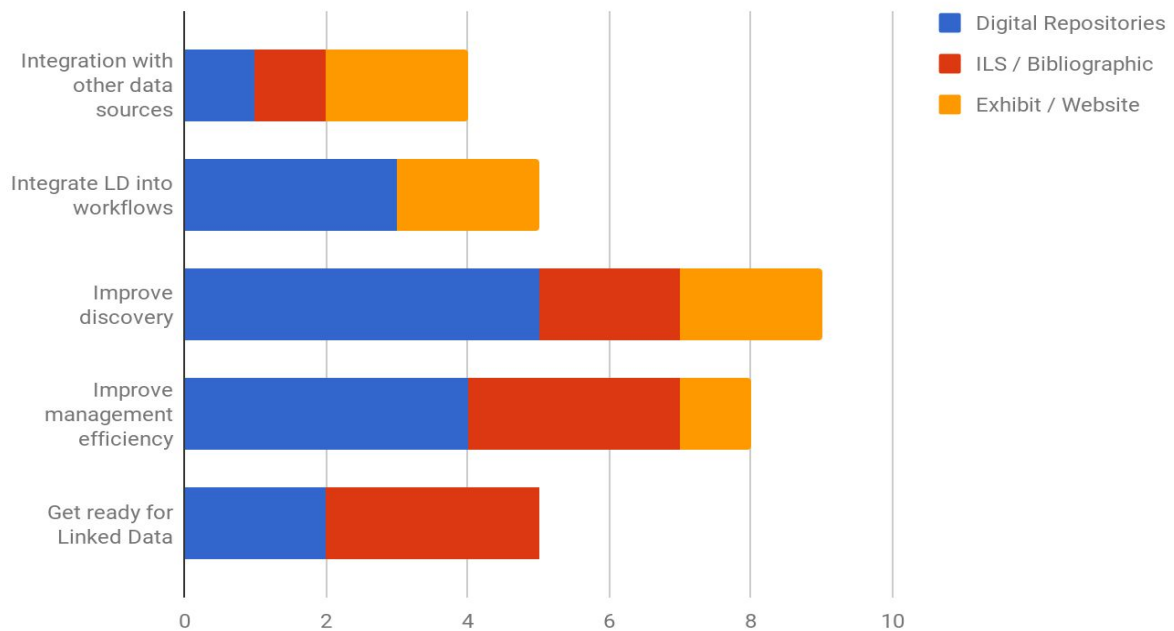
1. Your name *
2. Your campus email address *
3. Brief description of Linked Data work
4. Link(s) to public documentation, proposal, presentations, etc.
5. What stage is the work in?
 - Planning
 - Underway
 - Completed
 - Other:
6. What collaborators (departments, units, organizations, etc.) are (or have been) involved?
7. What are the goals of this effort? Was it designed to address any specific use case(s)?
8. What functional areas or existing systems are involved?
 - ILS
 - Digital repository
 - Archival management
 - Exhibit systems
 - Website
 - Non-library systems
 - Other:
9. What technologies or software are involved? (e.g. Alma, Aleph, Samvera, Fedora, Kuali, OpenRefine, SPARQL, etc.)
10. What kinds of metadata standards or authorities are involved?
11. What has been most challenging?
12. What has been most successful?
13. What conclusions, if any, have you drawn from the work?

The projects and activities in [Appendix 1](#) marked with an asterisk were submitted via the survey, and formed the basis for a preliminary analysis and were included in the UC Linked Data Project Team’s UCDFx presentation¹⁰. A sampling of that presentation is included below:

Areas of work by campus

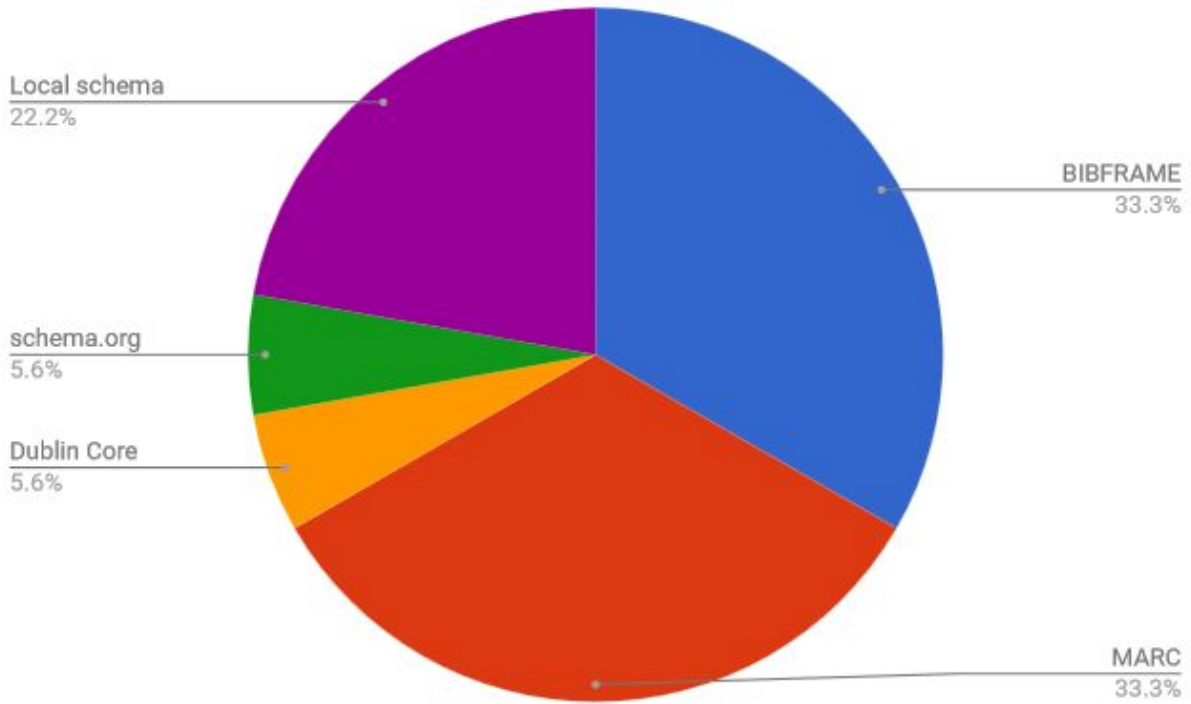


Goals

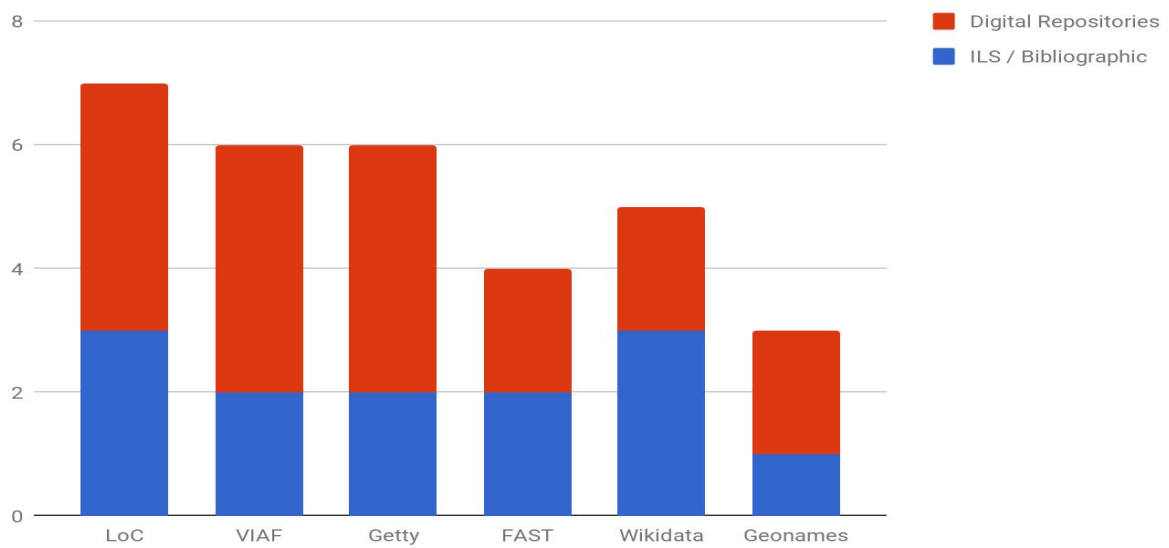


¹⁰ UC Linked Data Project Team Update, UCDFx, 2018
<https://docs.google.com/presentation/d/1qxHUz7vFkoIVc9YHLARxosgqz8ibja3KMI9IL0CsTio/edit?usp=sharing>

Metadata standards



Vocabularies



Appendix 3: Snapshots of External Linked Data Projects & Activities

Authority control, reconciliation, identify management

Cornell LD4L-Labs Questioning Authority gem

Three primary areas for investigation and development:

1. UI integration for lookup and use - currently Samvera and VitroLib
2. Authority normalization layer - standardizes and interprets query results from multiple linked data vocabularies
3. Authority data caching - comparison of options for caching authorities to address stability issues with direct access to externally hosted authorities

[https://wiki.duraspace.org/display/Id4L/LABS/Samvera+%28aka+Hydra%29+Community+Linked+Data+Support#Samvera\(akaHydra\)CommunityLinkedDataSupport-ApplicationUI](https://wiki.duraspace.org/display/Id4L/LABS/Samvera+%28aka+Hydra%29+Community+Linked+Data+Support#Samvera(akaHydra)CommunityLinkedDataSupport-ApplicationUI)

OCLCs Linked Data Wikibase prototype

Prototype service to reconcile legacy bibliographic information to linked data entities, create and mint new entities, and view, create, and edit relationships between entities. Also provided a Linked Data native cataloging interface for creating new and managing existing authorities as well as a prototype discovery interface and API for integration with external cataloging and discovery platforms.

<https://www.oclc.org/research/themes/data-science/linkedata/linked-data-prototype.html>

Western Name Authority File

Shared authority management for Mountain West Digital Library, currently in development:

https://www.ims.gov/sites/default/files/ig-72-16-0002_proposal_narrative.pdf

University of North Texas Names

System for creation of name authorities for 1) creation of locally relevant authorities that are not established in external vocabularies, and 2) records for aggregating external authority links. They are also using records from this system for creation of NACO submissions.

<http://digital2.library.unt.edu/name/>

Development and/or implementation of linked data based systems

VIVO, Vitro and VitroLib

VIVO is an open source linked data based software for research and scholarship portals. Vitro is the integrated ontology editor, for loading, and editing ontologies. VitroLib is a LD4P project to develop an application, like VIVO, but customized for original cataloging in RDF.

<https://wiki.duraspace.org/display/ld4LABS/The+VidroLib+Metadata+Editor>

Fedora 4

Fedora 4 is an open source, linked data platform server (W3C standard) and repository system. It includes "native" RDF expressions of object properties, uses URIs as names for things, HTTP URIs, and a RESTful API. Fedora 4 is a substantial

<https://wiki.duraspace.org/display/FF/Fedora+Repository+Home>

Hyrax Samvera Application

An open source, community developed, linked data based, digital repository application. Hyrax uses PCDM for modeling and Fedora 4 for data management.

<http://samvera.org/samvera-open-source-repository-framework/technology-stack/>

Islandora CLAW

An open source, community development project to create a new, linked data based version of the Islandora digital repository application. Islandora CLAW uses PCDM for modeling and Fedora 4 for data management.

<https://islandora.ca/CLAW>

Ex Libris

ILS vendor Ex Libris is adding linked data capabilities to their Alma and Primo systems.

<https://developers.exlibrisgroup.com/blog/Linked-Library-Data>

Publication of linked open data sets and vocabularies

Europeana

The EU digital platform for digital cultural heritage collections, in addition to aggregation and publication service, they enrich submitted metadata through reconciliation to linked data vocabularies. They also make their data available as linked data to be accessed and queried through their SPARQL API.

<https://pro.europeana.eu/page/linked-open-data>

British National Bibliography

The British Library began publishing a Linked Open Bibliography Data as a part of their data services in 2011. The bibliography data covers published and forthcoming books as well as serial publications. The service provides multiple accesses to the data set, including bulk downloads.

<http://www.bl.uk/bibliographic/download.html>

<http://bnb.data.bl.uk/>

Swedish Union catalog

The Swedish Union Catalog has been published as linked data since 2008. The highlight of this project is linking to external resources, in particular, DBpedia.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.181.860&rep=rep1&type=pdf>)

Controlled vocabularies

Many well-known established controlled vocabularies have been published as linked data. These vocabularies include LCSH, VIAF, AGROVOC, Getty vocabulary, Mesh, NDLSH.

- LC Linked Data Service
<https://id.loc.gov/>
- Virtual International Authority File (VIAF)
<https://viaf.org/>
- AGROVOC thesaurus
<http://aims.fao.org/standards/agrovoc/concept-scheme>
- National Diet Library List of Subject Headings (NDLSH)
<http://www.ndl.go.jp/en/data/ndla.html>
- Getty Vocabularies
<http://www.getty.edu/research/tools/vocabularies/>
- Medical Subject Headings (MeSH)
<https://www.nlm.nih.gov/mesh/>

Conversion of MARC data to linked data

Program for Cooperative Cataloging (PCC) Linked Data Strategies

- *ISNI Pilot*
Developing an understanding of ISNI tools and systems, creating PCC documentation and training for its members, and putting in place member supports for cost-effective scaling of broader PCC involvement in ISNI.
<https://wiki.duraspace.org/display/PCCISNI/PCC+ISNI+Pilot+Home>
- *Linked Data Advisory Committee report: LD infrastructure models*
Investigation into unresolved areas of linked data implementation and infrastructure models where the PCC should focus its resources.
<https://www.loc.gov/aba/pcc/documents/LinkedDataInfrastructureModels.pdf>
- *Task Group on URIs in MARC*
Charged with identifying unresolved issues with and developing best practices for including URIs in MARC bibliographic and authority records.
<https://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html>
- *Task Group on Identity Management in NACO*
Charged with defining the differences between “authority control” and “identity management,” investigating reasons for local authority creation, and identifying use cases where library authority data could be put to non-library uses.

<https://www.loc.gov/aba/pcc/documents/Identity-management-NACO-PCC-TG.pdf>

Stanford LD4P Tracer Bullets and Data-Pipeline projects

The Tracer Bullets project is focused on developing technical services workflows, including copy cataloging by acquisitions, original cataloging, deposit of single items, and batch deposit of collections of resources into the Stanford Digital Repository. The Data-Pipeline project is developing a continuous data transformation pipeline to convert MARC21 data created and managed in their ILS into BIBFRAME compliant RDF for use in other applications, as well as potentially serving as a prototype for other data pipelines.

<https://wiki.duraspace.org/display/LD4P/Stanford+Tracer+Bullets>

<https://github.com/sul-dlss-labs/ld4p-data-pipeline/wiki>

Linked Data for Production: Pathway to Implementation

Two year project that builds upon the work of Linked Data for Production (LD4P) Phase 1 and Linked Data for Libraries Labs (LD4L Labs), with the goal of beginning the implementation phase of the cataloging community's shift to linked data for the creation and manipulation of their metadata. The project is a collaboration between Cornell, Harvard, Stanford, the University of Iowa, and PCC and includes the goal of developing a larger cohort of libraries able to create and reuse linked data through a cloud-based sandbox editing environment.

SHARE-VDE Project

This project has undertaken large scale conversion, reconciliation, and enhancement of MARC21 data into the BIBFRAME data model. This converted data is forming the basis for the development of the Virtual Discovery Environment a prototype linked data based discovery tool.

<http://share-vde.org>

Discovery & visualization

BBC Linked Data Platform

The BBC's linked data Platform supplies information to many of their content focused (i.e. music, sports, news) tools and sites, both for internal management and external User Interface enhancement. They also make updates directly in linked data sources (Wikipedia, MusicBrainz), and have publically released their ontologies.

<http://www.bbc.co.uk/>

<https://www.bbc.co.uk/ontologies>

Libhub (Innovative and Zepheira)

Goal: "publish BIBFRAME resources to the Web, cross-link resources which are common among libraries, and, through cross-linking improve the ability for people to discover these resources on the open Web."

<https://libhub.squarespace.com/faq>

Library.Link network (Zepheira)

The Library.Link Network (LLN) is a data publishing platform that uses linked data standards and web technology on behalf of libraries to make their content available and accessible on the Web.

<http://library.link/>

The Advanced Research Consortium Big Diva project

BigDIVA (Big Data Infrastructure Visualization Application) is a graph based visualization and navigation interface for browsing, searching, and interacting with the ARC (Advanced Research Consortium) catalog.

<http://bigdiva.org/>

Appendix 4: Glossary of Terms¹¹

Actionable [Machine]: An object is said to be machine actionable when it is in a form that allows a computer to interact with it in some automated manner.

Application Programming Interface (API): An application programming interface (API) is a set of communication protocols that provide a clearly defined method of communication between various software components, programs, or network services.

BIBFRAME: Short for Bibliographic Framework—a data model created for bibliographic description. The design of BIBFRAME began in 2011 through a partnership between the Library of Congress and Zepheira. BIBFRAME’s goals include the replacement of MARC encoding standards with methods that integrate Linked Data principles in order to make bibliographic data more useful both within the library professional community and to the world at large.

Controlled Vocabulary: [Semantic Web community definition from W3C] Carefully selected sets of terms that are used to describe units of information; used to create taxonomies, thesauri and ontologies. In traditional settings the terms in the controlled vocabularies are words or phrases, in a linked data setting then they are normally assigned unique identifiers (URIs) which in turn link to descriptive phrases.

Disambiguate: A process directed at distinguishing between distinct entities. Graph: A graph is data arrangement that consist of nodes (objects) connected to each other via edges (relationships). A family tree is a common example of a graph where the persons represents nodes (John, Jane, etc.) and relationships represent edges (child, parent, etc.).

Entity: A thing with a distinct identity, such as a person, corporation, place, object, etc. In the sense of an entity-attribute-value model, an entity is synonymous with the Subject of an RDF Triple.

¹¹ Full list of terms available at <https://github.com/ucsdlb/uc-linked-data/wiki/Notable-Definitions>

Five Star Linked Open Data: An incremental framework for deploying data. Tim Berners-Lee, the inventor of the Web and initiator of the Linked Data project, suggested a 5 star deployment scheme for Linked Open Data. The 5 Star Linked Data system is cumulative. Each additional star presumes the data meets the criteria of the previous step(s). 5 Star Linked Open Data includes an Open License (expression of rights) and assumes publications on the public Web. Organizations may elect to publish 5 Star Linked Data, without the word "open", implying that the data does not include an Open License (expression of rights) and does not imply publication on the public Web. The 5 stars are: 1) Publish data on the Web in any format (e.g., PDF, JPEG) accompanied by an explicit Open License (expression of rights); 2) Publish structured data on the Web in a machine-readable format (e.g., XML); 3) Publish structured data on the Web in a documented, non-proprietary data format (e.g., CSV, KML); 4) Publish structured data on the Web as RDF (eg Turtle, RDFa, JSON-LD, SPARQL); and 5) In your RDF, have the identifiers be links (URLs) to useful data sources.

Graph [database]: A collection of objects (represented by "nodes") any of which may be connected by links between them. This contrasts with relational databases that, with the aid of relational database management systems, permit managing the data without imposing implementation aspects like physical record chains. Graph databases, by design, allow simple and fast retrieval of complex hierarchical structures that are difficult to model in relational systems.

Inference: Inference is the process of deriving logical conclusions from a set of starting assumptions. Using Linked Data, existing relationships are modeled as a set of (named) relationships between resources. Linked Data helps humans and machines to find new relationships through automatic procedures that generate new relationships based on the data and based on some additional information in the form of a vocabulary.

Linked Data: According to the W3C the term Linked Data refers to a set of best practices for publishing structured data on the Web that includes the use Uniform Resource Identifiers (URIs) as names for things, the use of HTTP URIs so that people can look up those names, insuring that when someone looks up a URI, provide useful information, and including links to other URIs so that users can discover more things. Additionally, Linked Data describes a semantic data structure based on collections of n-triples preferably (but not necessarily) serialized as RDF. See <https://www.w3.org/wiki/LinkedData>.

Linked Data API: A REST API that allows data publishers to provide URLs to lists of things and clients to retrieve machine-readable data from those URLs.

Linked Data Platform: A specification that defines a REST API to read and write Linked Data for the purposes of enterprise application integration. The Linked Data Platform describes the use of a REST API for accessing, updating, creating and deleting resources from servers.

Machine Readable Data: Data formats that may be readily parsed by computer programs without access to proprietary libraries. For example, CSV, TSV and RDF formats are machine readable, but PDF and Microsoft Excel are not. Creating and publishing data following Linked Data principles helps search engines and humans to find, access and re-use data. Once information is found, computer programs can re-use data without the need for custom scripts to manipulate the content.

Persistent Identifier Scheme: A mechanism for resolution of virtual resources. Persistent Uniform Resource Locator (PURLs) implement one form of persistent identifier for virtual resources. PURLs are valid URLs and their components must map to the URL specification. The scheme part tells a computer program, such as a Web browser, which protocol to use when resolving the address. The scheme used for PURLs is generally HTTP. Other persistent identifier schemes include Digital Object Identifiers (DOIs), Life Sciences Identifiers (LSIDs) and INFO URIs. All persistent identification schemes provide unique identifiers for (possibly changing) virtual resources, but not all schemes provide curation opportunities.

Reconciliation: There are several kinds of reconciliation, most frequently in discussions of linked data and library metadata it refers to the process of matching text terms (names, places, etc.) from linked data sources and replacing them in metadata records with URIs or storing the URIs, and so disambiguating from other identities that may have the same or similar string names (e.g. Taj Mahal the building or the musician). URI reconciliation is used to describe a process of aligning/grouping synonymous URIs (URIs that refer the same entity).

RDFa: Resource Description Framework in Attributes (RDFa) is a syntax encoded in HTML documents. RDFa allows web developers to imbed RDF triples in webpages using HTML attributes. This provides a set of markup attributes to augment the visual information on the Web with machine-readable hints. It is a standard of the World Wide Web Consortium.

Semantic Web: An evolution or part of the World Wide Web that consists of machine-readable data in RDF and an ability to query that information in standard ways (e.g. via SPARQL).

SPARQL: SPARQL Protocol and RDF Query Language (SPARQL) defines a query language for RDF data, analogous to the Structured Query Language (SQL) for relational databases. A family of standards of the World Wide Web Consortium.

Triple: An RDF statement, consisting of two things (a "subject" and an "object") and a relationship between them (a verb, or "predicate"). This subject-predicate-object triple forms the smallest possible RDF graph (although most RDF graphs consist of many such statements).

Triple Store: A colloquial phrase for an RDF database that stores RDF triples.

Uniform Resource Identifier (URI): In information technology, a Uniform Resource Identifier (URI) is a string of characters used to identify a resource. Such identification enables interaction with representations of the resource over a network using specific protocols. In practical terms, to human readers look like the URLs used to navigate the World Wide Web. URIs, however, by convention, are intended to be permanent identifiers for a resource, regardless of it might live (or move to) on the network. In other words, an item's URL could change, if, for example, a web-based resource moved to another hosting environment, but its URI would not and any person or machine that traverses the URI would be directed to the current URL for the resource.

Vocabulary Alignment: The process of analyzing multiple vocabularies to determine terms that are common across them and to record those relationships. Tightly coupled with the process of URI Reconciliation.

Appendix 5: Bibliography / List of Resources

- Aalberg, Trond, et al. "Benchmarking and Evaluating the Interpretation of Bibliographic Records." *International Journal on Digital Libraries*, 2018, pp. 1–23, doi:<https://doi.org/10.1007/s00799-018-0233-2>.
- Advanced Research Consortium at Texas A&M U. *BigDIVA*. 2015, <http://bigdiva.org/>.
- Alemu, Getaneh, et al. "Linked Data for Libraries: Benefits of a Conceptual Shift from Library-specific Record Structures to RDF-based Data Models." *New Library World*, vol. 113, no. 11/12, 2012, pp. 549–70, <https://www.emeraldinsight.com/doi/full/10.1108/03074801211282920>.
- "Background Paper." *IMLS Shareable Authorities Forum*, <https://confluence.cornell.edu/display/sharedauth/Background+Paper>. Accessed 24 Apr. 2018.
- Bensmann, Felix, et al. "Interlinking Large-Scale Library Data with Authority Records." *Frontiers in Digital Humanities*, vol. 4, 2017, doi:<https://doi.org/10.3389/fdigh.2017.00005>.
- Biagetti, Maria Teresa, et al. "Quality Evaluation of Online Library Catalogues, Advanced Discovery Tools and Linked Data Technologies." *The Evaluation of Research in Social Sciences and Humanities*, edited by Andrea Bonaccorsi, Springer, 2018, pp. 185–209.
- Boehr, Diane L., and B. Bushman. "Preparing for the Future: National Library of Medicine's Project to Add MeSH RDF URIs to Its Bibliographic and Authority Records." *Cataloging & Classification Quarterly*, vol. 56, no. 2–3, Feb. 2018, pp. 262–72, doi:[10.1080/01639374.2017.1382642](https://doi.org/10.1080/01639374.2017.1382642).
- Byrne, Gillian, and Lisa Goddard. "The Strongest Link: Libraries and Linked Data." *D-Lib Magazine*, vol. 16, no. 11/12, 2010, <http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/november10/byrne/11byrne.print.html>.
- Candela, Gustavo, et al. "Migration of a Library Catalogue into RDA Linked Open Data." *Semantic Web Journal*, vol. to appear, <http://www.semantic-web-journal.net/system/files/swj1453.pdf>.
- Casalini, Michele, et al. *National Strategy for Shareable Local Name Authorities National Forum : White Paper*. report, 29 Mar. 2018, <http://ecommons.cornell.edu/handle/1813/56343>.
- Clark, Jason A., and Scott W. H. Young. "Linked Data Is People: Building a Knowledge Graph to Reshape the Library Staff Directory." *Code4Lib Journal*, vol. 36, 2017, <http://journal.code4lib.org/articles/12320>.
- Corcho, Oscar, et al. "Ontology Engineering in the Era of Linked Data." *Bulletin of the Association for Information Science and Technology*, vol. 41, no. 4, 2015, <https://www.asist.org/publications/bulletin/apr-15/ontology-engineering-in-the-era-of-linked-data/>

- Crowe, Katherine, and Kevin Clair. "Developing a Tool for Publishing Linked Local Authority Data." *Journal of Library Metadata*, vol. 15, no. 3–4, Oct. 2015, pp. 227–40, doi:[10.1080/19386389.2015.1099993](https://doi.org/10.1080/19386389.2015.1099993).
- Deliot, Corine, et al. "The British National Bibliography: Who Uses Our Linked Data?" *Proceedings of International Conference on Dublin Core and Metadata Applications*, 2016, pp. 24–33, <http://dcpapers.dublincore.org/pubs/article/viewFile/3820/2005>.
- El-Sherbini, Magda. *Improving Resource Discoverability for Non-Roman Language Collections*. 2017, <http://library.ifla.org/1982/>.
- Färber, Michael, et al. "Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO." *Semantic Web*, vol. 9, no. 1, 2018, pp. 77–129.
- Godby, Carol Jean, et al. *Library Linked Data in the Cloud : OCLC's Experiments with New Models of Resource Description*. Morgan & Claypool, 2015.
- Godby, Carol Jean, and Karen Smith-Yoshimura. "From Records to Things: Managing the Transition from Legacy Library Metadata to Linked Data." *Bulletin of the Association for Information Science and Technology*, vol. 43, no. 2, Jan. 2017, pp. 18–23, doi:[10.1002/bul2.2017.1720430209](https://doi.org/10.1002/bul2.2017.1720430209).
- Gonzales, Brigid M. "Linking Libraries to the Web: Linked Data and the Future of the Bibliographic Record." *Information Technology and Libraries*, vol. 33, no. 4, 2014, pp. 10–22, doi:<https://doi.org/10.6017/ital.v33i4.5631>.
- Hallo, Maria, et al. "Current State of Linked Data in Digital Libraries." *Journal of Information Science*, vol. 42, no. 2, 2016, pp. 117–27, <http://journals.sagepub.com/doi/abs/10.1177/0165551515594729>.
- Hannemann, Jan, and Jürgen Kett. *Linked Data for Libraries*. 2010, <https://www.ifla.org/past-wlic/2010/149-hannemann-en.pdf>.
- Hastings, Robin. "Linked Data in Libraries: Status and Future Direction." *Computers in Libraries*, vol. 35, no. 9, 2015, <http://www.infotoday.com/cilmag/nov15/Hastings--Linked-Data-in-Libraries.shtml>.
- Henneken, Edwin. "Unlocking and Sharing Data in Astronomy." *Bulletin of the Association for Information Science and Technology*, vol. 41, no. 4, 2015, <https://www.asist.org/publications/bulletin/apr-15/rdap-review/>.
- Hooland, Seth van, et al. "Evaluating the Success of Vocabulary Reconciliation for Cultural Heritage Collections." *Journal of the Association for Information Science and Technology*, vol. 64, no. 3, 2013, pp. 464–479, doi:<https://doi.org/10.1002/asi.22763>.
- Hooland, Seth van, and Ruben Verborgh. *Linked Data for Libraries, Archives, and Museums*. Facet Publishing, 2014.
- "IMLS Focus Summary Report: National Digital Platform." *Institute of Museum and Library Services*, 23 July 2015,

<https://www.ims.gov/publications/ims-focus-summary-report-national-digital-platform>.

Isaac, Antoine, and Thomas Baker. "Linked Data Practice at Different Levels of Semantic Precision: The Perspective of Libraries, Archives and Museums." *Bulletin of the Association for Information Science and Technology*, vol. 41, no. 4, 2015,

<https://www.asist.org/publications/bulletin/apr-15/linked-data-practice-at-different-levels-of-semantic-precision/>.

Kroeger, Angela. "The Road to BIBFRAME: The Evolution of the Idea of Bibliographic Transition into a Post-MARC Future." *Cataloging & Classification Quarterly*, vol. 51, no. 8, 2013,

<https://www.tandfonline.com/doi/abs/10.1080/01639374.2013.823584>.

Kumar, Vinit. "A Model for Content Enrichment of Institutional Repositories Using Linked Data." *Journal of Web Librarianship*, vol. 12, no. 1, 2018, pp. 46–62.

"Linked Data Wikibase Prototype." *OCLC Research*, 29 Apr. 2018,

<https://www.oclc.org/research/themes/data-science/linkeddata/linked-data-prototype.html>.

"Linked Open Data." *Europeana*, <https://pro.europeana.eu/page/linked-open-data>.

The Library.Link Network. <http://library.link/>. Accessed 29 Apr. 2018.

MacEwan, Andrew. *ISNI and VIAF: Authority Files and Identity*.

<https://www.oclc.org/content/dam/oclc/events/2016/IFLA2016/presentations/ISNI-and-VIAF-Authority-Files-and-Identity-Management.pdf>. Authority Data on the Web, Dublin, OH.

Malmsten, Martin. *Exposing Library Data as Linked Data*. 2009.

Metadata Policy Project Team. "University of California Libraries Metadata Sharing Policy." *UC Libraries*, 2015,

https://libraries.universityofcalifornia.edu/groups/files/sag2/docs/Final_UC_Libraries_Metadata_Sharing_Policy_May8_2015.pdf.

Miller, Eric, and U. Uch. Ogbuji. "Linked Data Design for the Visible Library." *Bulletin of the Association for Information Science and Technology*, vol. 41, no. 4, 2015,

<https://www.asist.org/publications/bulletin/apr-15/linked-data-design-for-the-visible-library/>

Mitchell, Erik. "A Linked Data Landscape : Critical Decisions for Data Licensing, Shared Standards, and System Design." *American Libraries Magazine*, 2016,

<https://americanlibrariesmagazine.org/2016/01/04/linked-data-landscape/>.

Mitchell, Erik T. "Building Blocks of Linked Open Data in Libraries." *Library Technology Reports*, vol. 49, no. 5, 2013.

---. "Library Linked Data : Research and Adoption." *Library Technology Reports*, vol. 49, no. 5, July 2013.

Morgan, Eric Lease, and LiAM. *Linked Archival Metadata: A Guidebook*. Version 0.99, 2014, <http://sites.tufts.edu/liam/>.

Page, Kevin R., et al. "Realising a Layered Digital Library : Exploration and Analysis of the Live Music Archive through Linked Data." *Proceedings of the 17th ACM/IEEE Joint Conference on Digital Libraries*, 2017, <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7991563>.

Papadakis, Ioannis, et al. "Linked Data URIs and Libraries: The Story So Far." *D-Lib Magazine*, vol. 21, no. 5/6, 2015, <http://www.dlib.org/dlib/may15/papadakis/05papadakis.html>.

PCC Linked Data Advisory Committee. *Linked Data Infrastructure Models: Areas of Focus for PCC Strategies*. 2017, <https://www.loc.gov/aba/pcc/documents/LinkedDataInfrastructureModels.pdf>.

PCC Steering Committee. *Charge for PCC Task Group on Identity Management in NACO*. 31 Mar. 2016, <https://www.loc.gov/aba/pcc/documents/Identity-management-NACO-PCC-TG.pdf>.

Program for Cooperative Cataloging. "PCC ISNI Pilot." *DuraSpace Wiki*, <https://wiki.duraspace.org/display/PCCISNI/PCC+ISNI+Pilot+Home>. Accessed 29 Apr. 2018.

Program for Cooperative Cataloging (Library of Congress). "PCC Task Group on URIs in MARC." *The Library of Congress*, <https://www.loc.gov/aba/pcc/bibframe/TaskGroups/URI-TaskGroup.html>. Accessed 29 Apr. 2018.

Sande, Miel Vander, et al. "Toward Sustainable Publishing and Querying of Distributed Linked Data Archives." *Journal of Documentation*, vol. 74, no. 1, 2018.

Schiavone, Luisa, et al. "Library Data Integration: The CoBiS Linked Open Data Project and Portal." *Digital Libraries and Multimedia Archives. IRCDL 2018. Communications in Computer and Information Science*, edited by Giuseppe Serra and Carlo Tasso, vol. 806, Springer, 2018.

Schilling, Virginia. "Transforming Library Metadata into Linked Library Data: Introduction and Review of Linked Data for the Library Community, 2003–2011." *Association for Library Collections & Technical Services*, 25 Sept. 2012, <http://www.ala.org/alcts/resources/org/cat/research/linked-data>.

Schmedding, Florian. "Incremental SPARQL Evaluation for Query Answering on Linked Data." *Proceedings of the Second International Conference on Consuming Linked Data*, vol. 782, pp. 49–60, http://ceur-ws.org/Vol-782/Schmedding_COLD2011.pdf.

- “SHARE-VDE Project.” *SHARE Virtual Discovery Environment*,
<http://share-vde.org/sharevde/clusters?l=en>. Accessed 29 Apr. 2018.
- Smith, MacKenzie, et al. *BIBFLOW: A Roadmap for Library Linked Data Transition*. 14 Mar. 2017,
https://bibflow.library.ucdavis.edu/wp-content/uploads/2017/03/bibflow_roadmap_revised_3_14_2017.pdf.
- Stvilia, Besiki, et al. “Toward Engaging Researchers in Research Identity Data Curation.” *Institute of Museum and Library Services, Award #LG-73-16-0006-16*, 2016,
<https://www.ims.gov/grants/awarded/lg-73-16-0006-16>.
- Taniguchi, Shoichi. “Is BIBFRAME 2.0 a Suitable Schema for Exchanging and Sharing Diverse Descriptive Metadata about Bibliographic Resources?” *Cataloging & Classification Quarterly*, vol. 56, no. 1, 2017, pp. 40–61, <https://doi.org/10.1080/01639374.2017.1382643>.
- UC Libraries Council of University Librarians (CoUL). *Charge: Working Group for the Systemwide ILS Planning Project*. 21 Nov. 2017,
https://libraries.universityofcalifornia.edu/groups/files/coul/docs/CoUL_SystemwideILS_Working_GroupCharge.pdf.
- UC Libraries DAMS Technology Project Team. *UC Libraries DAMS Technology Report: Assessment of a Long-Term Solution for the UC Libraries Systemwide DAMS*. 18 Nov. 2016,
https://wiki.library.ucsf.edu/display/UCLDTP/UC+Libraries+DAMS+Technology+Project+Home?preview=/380146695/398659887/DOC_DAMS_FinalReport_18Nov2016.docx#UCLibrariesDAMSTechnologyProjectHome-FinalReport.
- UC Libraries Direction & Oversight Committee. *Linked Data Project Team Charge*. Sept. 2017,
<https://libraries.universityofcalifornia.edu/groups/files/doc/docs/Linked%20Data%20Charge%20-%20Final%202017.pdf>.
- University of Utah. “Linking People: Developing Collaborative Regional Vocabularies.” *Institute of Museum and Library Services, Award #LG-72-16-0002-16*, 2016,
<https://www.ims.gov/grants/awarded/lg-72-16-0002-16>.
- “VIAF: Overview.” *OCLC*, <https://www.oclc.org/en/viaf.html>.
- VIAF: *The Virtual International Authority File*. viaf.org/.
- “Vision and Priorities.” *UC Libraries*,
<https://libraries.universityofcalifornia.edu/about/vision-and-priorities>. Accessed 29 Apr. 2018.
- W3C Incubator Group. *Library Linked Data Incubator Group: Datasets, Value Vocabularies, and Metadata Element Sets*. 25 Oct. 2011,
<http://www.w3.org/2005/Incubator/lld/XGR-lld-vocabdataset-20111025/>.

---. *Library Linked Data Incubator Group Final Report*. 25 Oct. 2011,
<https://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>.

---. *Library Linked Data Incubator Group: Use Cases*. 25 Oct. 2011,
<https://www.w3.org/2005/Incubator/ld/XGR-ld-usecase-20111025/>.

Wenige, Lisa, and Johannes Ruhland. "Retrieval by Recommendation: Using LOD Technologies to Improve Digital Library Search." *International Journal on Digital Libraries*, 2017,
<https://doi.org/10.1007/s00799-017-0224-8>.